

Добрынина Ксения Сергеевна

О МЕТОДИКЕ РАБОТЫ НАД КОНКОРДАНСАМИ

Статья раскрывает основное понятие "конкорданс", историю его возникновения, структуру, основные функции и задачи. Проводится противопоставление между словарями и конкордансами по нескольким шкалам. Уделяется особое внимание многочисленным и разнообразным по содержанию и принципам построения словарям языка художественных авторов, а также подробно рассматривается структура словарной статьи открытого конкорданса У. Шекспира.

Адрес статьи: www.gramota.net/materials/2/2012/1/15.html

Источник

Филологические науки. Вопросы теории и практики

Тамбов: Грамота, 2012. № 1 (12). С. 59-63. ISSN 1997-2911.

Адрес журнала: www.gramota.net/editions/2.html

Содержание данного номера журнала: www.gramota.net/materials/2/2012/1/

© Издательство "Грамота"

Информацию о том, как опубликовать статью в журнале, можно получить на Интернет сайте издательства: www.gramota.net
Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: voprosy_phil@gramota.net

они могут развиваться по ходу сюжета в силу того, что действующее лицо априори представляет собой определённый набор признаков, то есть является многосоставным образованием, легко распадающимся и столь же легко комплектующимся в новое.

Список литературы

1. **Литературный энциклопедический словарь** / В. М. Коженикова, П. А. Николаева. М.: Советская энциклопедия, 1987. 752 с.
2. **Пропп В. Я.** Морфология «волшебной сказки». Исторические корни волшебной сказки. М.: Лабиринт, 1998. 512 с.
3. **Фрейдберг О. М.** Античные теории языка и стиля. М. - Л.: Гослитиздат, 1936. 440 с.
4. **Rowling J. K.** Harry Potter and the Order of Phoenix. London: Bloomsbury Publishing, Plc., 2003. 956 p.
5. **Rowling J. K.** Harry Potter and the Philosopher's Stone. London: Bloomsbury Publishing, Plc., 1997. 223 p.

AUTHOR'S FAIRY TALE CHARACTERS SYSTEM

Irina Ivanovna Grishina, Ph. D. in Philology
 Department of Foreign Languages
 Siberian Federal University
 pkpel@yandex.ru

An author's fairy tale characters and personages system is considered in the article in connection with fairy-tale plot complication. This system can be corrected by the author both quantitatively and qualitatively.

Key words and phrases: character; fairy tale; characters levels; character semantic features; personage.

УДК 81-139

Статья раскрывает основное понятие «конкорданс», историю его возникновения, структуру, основные функции и задачи. Проводится противопоставление между словарями и конкордансами по нескольким шкалам. Уделяется особое внимание многочисленным и разнообразным по содержанию и принципам построения словарям языка художественных авторов, а также подробно рассматривается структура словарной статьи открытого конкорданса У. Шекспира.

Ключевые слова и фразы: словарь; конкорданс; лексикография; корпус языка; словарная статья.

Ксения Сергеевна Добрынина

Кафедра теории и практики перевода

Оренбургский государственный педагогический университет

Ks.thekind@yandex.ru

О МЕТОДИКЕ РАБОТЫ НАД КОНКОРДАНСАМИ[©]

Что такое *конкорданс*? Довольно полное, на наш взгляд, определение конкордансу было дано А. С. Гердом еще в начале 1980-х гг. [3]. Это «упорядоченный список словоформ (или слов) с указанием всех вхождений того или иного слова в заданный массив текстов», т.е. речь идет о полном, законченном тексте или их группе, при этом для каждой словоформы «приведен контекст, иллюстрирующий данное словоупотребление», что отличает конкорданс от обычных словоуказателей. «Анализируемый массив текстов разбивается на словоформы с определенным порядком следования, сопровождаемые статистической информацией и контекстом» (это одно из кратких описаний методики составления конкордансов, относящееся еще к середине 1980-х) [2; 5].

Другими словами, конкорданс - это традиционный, давно известный, но до сих пор недостаточно оцененный способ изучения текста. Он дает полный индекс слов в ближайших и расширенных контекстах. Компьютерный конкорданс позволяет сравнить все контексты употребления слова, проанализировать их, увидеть слово в самом тексте художественного произведения. Конкорданс представляет собой один из эффективных инструментов изучения литературного текста.

К примеру, словарная статья открытого конкорданса по произведениям У. Шекспира (*OpenSourceShakespeare.com*) выглядит так [7]:

JEALOUS occurs 51 times in 45 lines within 21 works

#	Work	Character	Line	Text
1	Much Ado about Nothing	Beatrice	671	The count is neither sad, nor sick, nor merry, nor well; but civil count, civil as an orange, and something of that <i>jealous</i> complexion.

Этот конкорданс называют *more than just a collection of texts*, он представляет собой «нечто большее, чем просто собрание текстов». Обратимся к истории конкордансов как отдельных типов словарей.

Английская национальная лексикография имеет давние традиции. Она формировалась в течение многих веков под влиянием различных тенденций, исторически сложившихся в западноевропейских странах, где роль кодификатора нормы выполняли академии [5]. Значительную роль в ее становлении сыграли специальные словари английского языка, а именно: учебные, словари синонимов, правильного словоупотребления и во многом писательские справочники к сочинениям Чосера, Мильтона, Шекспира, а также конкордансы к Библии, история создания которых восходит к XVI в. [6]. Первый конкорданс, к Вульгате, был создан в 13 веке Хьюго де Сент-Шером (Hugues de Saint-Cher), которому помогали 500 монахов. Именно из-за объема работы по составлению конкордансов в докомпьютерные времена конкордансы были созданы только для трудов высочайшей важности - Вед, Библии, Корана или произведений У. Шекспира [1].

Следует отметить важность гипотезы Л. П. Ступина [12] о теории литературного авторитета, ставшей основой для выделения в английской лексикографии словарей *предписывающего* типа, составители которых рекомендовали соотечественникам норму писательского словоупотребления, стремясь сделать родной язык более «элегантным» [9]. Хочется отметить, что вообще заметное влияние на характер языка имеют различные общественные течения и взгляды. Так, например, после революции в язык хлынул широкий поток различных жаргонизмов, арготизмов и провинциализмов. Эти слои лексики проникли и в художественную литературу.

Многие выдающиеся писатели, драматурги, артисты сыграли важную роль в развитии того или иного литературного языка. Такова, например, роль Пушкина и целой плеяды классиков русской литературы в России, роль Данте в Италии, Сервантеса в Испании, Чосера и Шекспира в Англии и т.д. Таким образом, конкордансы и глоссарии ко всему творчеству и отдельным произведениям известных поэтов и прозаиков стали источником готового иллюстративного материала для словарных статей первых толковых словарей английского литературного языка, так как имели в своем распоряжении полные своды цитат из отдельных произведений или всех сочинений писателей [6; 22].

Именно словари языка выдающихся поэтов и прозаиков Великобритании становятся надежным источником не только вербального иллюстративного материала для микроструктуры толковых словарей, но и других информационных категорий, которые постепенно вошли в мега-, макро- и микроструктуру академических толковых словарей английского литературного языка [6].

Ключевыми для данной статьи являются понятия **конкорданса**, **корпуса**. Поиск в корпусе данных позволяет по любому слову построить конкорданс - список всех употреблений данного слова в контексте со ссылками на источник.

Анализ современных лексикографических продуктов и аналитических работ по авторской и - шире - общей лексикографии показывает, что термин «словарь» представляет собой родовое обозначение, соответствующее общезыковому употреблению слова. Фактически, в широком смысле слова, это некоторым образом упорядоченный список символов, выбранный из некоторого множества текстов.

Словарь в широком смысле подразумевает определенный способ выборки словника, множество текстов, из которого делается эта выборка, характеристику вокабулы, дополнительную информацию, которая сопровождает каждую вокабулу, возможности филиации значений и их толкований. Таким образом, ключевые различия проявляются как на уровне микроструктуры - того, что представляет собой единица описания плюс структура словарной статьи, - так и на уровне макроструктуры - пространства текстов, являющегося базой для составления словаря. На основе описания этих параметров можно говорить и о принципиальном различии между конкордансами и собственно словарями (в узком смысле слова); оно состоит в принципах отбора, подачи и описания лексических единиц, в постановке различных целей описания и следовании им.

Минималистское определение конкорданса с необходимостью включает представление о корпусе (наборе текстов, отобранных и препарированных с определенной целью). Тогда **конкорданс** к некоторому корпусу - это список словоупотреблений (элементов корпуса) с отсылкой ко всем контекстам. Противопоставление между словарями и конкордансами идет сразу по нескольким шкалам - репрезентативности, ориентации на инвариант, смыслового или грамматического анализа. Можно сформулировать следующие противопоставления:

- Словарный подход к описанию лексики ориентирован на репрезентативность и тем самым на **нормативность**, а корпусный вариант (конкорданс) - на **исчерпывающее** описание.

- Словарь, анализируя различные употребления лексемы в разных значениях, стремится к нахождению **инварианта**, конкорданс - к **вариативности** и ставит своей первой задачей отразить все случаи употребления слова. Поэтому представление слова в конкордансе ставит во главу угла примеры (контексты словоупотреблений), а в словаре - словарную статью.

- На различном понимании термина **полноты** базируется еще одно принципиальное противопоставление словаря и конкорданса: полнота словаря определяется стремлением к исчерпывающему описанию значений, полнота конкорданса определяется исчерпывающим характером описания соответствующего корпуса.

- Из этого вытекает принципиальная необходимость грамматической (морфологической) информации в конкордансе, что помогает охарактеризовать и - если нужно - различить формы, и необходимость семантического анализа (описания или толкования значений) в словаре.

Итак, можно сформулировать представление об «идеальном» (прототипическом) словаре и прототипическом конкордансе. Словарный подход стремится выделить некоторую сущность и истолковать ее. Прототипический словарь - это нормативный толковый словарь, содержащий большое количество семантической

информации, в частности, семантических и стилистических помет. Прототипический конкорданс обладает полнотой тезаурусного типа (словаря, в котором максимально полно представлены слова языка с примерами их употребления в тексте, что в полном объеме осуществимо, да и то с оговорками, лишь для мертвых языков), он является первой производной корпуса, «нарезанным» корпусом, так что в идеале совокупность примеров всех употреблений составляют корпус. Он обязан быть полным в отношении учета абсолютно всех словоупотреблений. От списка слов (словника) его отличает наличие морфологической характеристики словоупотреблений. В конкордансе морфологическое описание реализуется как лемматизация, позволяющая приписать каждой словоформе граммемную характеристику. Это, в свою очередь, позволяет различить омонимичные формы. Реальной такая постановка задачи и стремление хоть сколько-нибудь приблизиться к данному прототипу возможны только в электронной форме сетевого ресурса.

Таким образом, различие между конкордансом и толковым словарем заключается в том, что конкорданс не предполагает установления структуры (филиации) значений регистрируемых слов и не обязательно включает толкование этих значений. Его полнота, в частности, требует, чтобы он был словарем регистрирующего типа. Задача Конкорданса состоит в том, чтобы отразить все возможные различия и особенности употреблений, так как заранее неизвестно, какие из них могут оказаться значащими и значимыми. В этом отношении Конкорданс представляет собой базу для созданий различного рода **дифференциальных**, в том числе и **толковых**, словарей [10].

Двуязычные конкордансы - это конкордансы, основанные на параллельных текстах. Тематическими конкордансами называют списки тем, которые охватывает книга, с содержанием сути этих тем. Наиболее известный из таких конкордансов - *Nave's Topical Bible* [23].

Конкордансы используются для решения следующих лингвистических задач:

- сравнение различных использований одного слова;
- анализ ключевых слов;
- анализ частотности слов и словосочетаний;
- поиск и исследование фраз и идиом;
- поиск перевода, например, терминологии;
- создание списков слов (что используется при публикации) [1].

Примеры конкордансов:

- Словарь-конкорданс публицистики Ф. М. Достоевского [11].
- *Nave's Topical Bible* [22].
- *AntConc* [20].
- *GlossaNet* [19].
- *Unitex* [21].

Вернемся к одной из целей данной статьи, а именно, рассмотреть словарную статью открытого конкорданса Шекспира. Представляется немаловажным заметить, что целую эпоху в английской писательской лексикографии составляют многочисленные и разнообразные по содержанию и принципам построения словари языка У. Шекспира [13-15; 17]. Первые из них были дифференциальными справочниками, обрабатывающими «трудные места» из произведений писателя. Лексикографическая обработка слов проводилась как посредством иллюстрации контекстами в конкордансах, так и с помощью толкований в глоссариях и толковых словарях.

Вершиной в составлении всех справочников к творчеству великого классика считается словарь Шмидта [24]. Автор осуществил на практике не только полноту словника, но и полноту разработки значений. Этот справочник явился исчерпывающим, что отличает его от всех предыдущих трудов лексикографов-составителей различных писательских словарей [8].

В XX в. вновь в больших количествах выходят конкордансы, глоссарии и толковые словари к Библии, Чосеру, Шекспиру [16; 18; 27]. Тем не менее, популярными словарями ко всему творчеству и к отдельным произведениям Шекспира в этот период времени по праву считаются конкордансы М. Спевака [25; 26]. Судя по словарям языка Шекспира, можно сделать вывод о том, что в XX столетии словари языка писателей разделились на две категории: полные и дифференциальные [8].

У представляемого здесь конкорданса - две основные функции: Во-первых, *поисковая функция*, - читатель может быстро найти фрагмент текста, даже если он запомнил только одно слово (например, *thy* или *Shadow*). Конкорданс полезен, если нужно сверить или подобрать цитату, уточнить текст оригинала. Это может потребоваться и студенту, и преподавателю, и составителю хрестоматии. Во-вторых, *эвристическая функция*, - она отличает конкорданс от именных указателей, так как в конкордансе при слове приведены также и контексты. Контексты могут дать новое понимание текста, служить хорошим подсказчиком и советчиком вдумчивому читателю. Обратимся к структуре словаря.

Открытый конкорданс является конкордансом словоформ ко всем поэтическим текстам произведений У. Шекспира, включая все пьесы, сонеты и стихотворения и содержит все словоформы, встречающиеся в 37 пьесах (комедиях, трагедиях и исторических хрониках), 154 сонетах и 5 поэмах.

Словарная статья включает следующие основные зоны:

- 1) Заголовочное слово.
- 2) Суммарная частота по всем текстам (возможно разбиение по типам/жанрам).
- 3) Примеры употребления с адресами и гиперссылками.

Указывается абсолютная частота лексемы, т.е. количество всех ее употреблений во всех рассматриваемых текстах. При необходимости дается разбиение по основным типам/жанрам текстов.

Например, если в строке поиска мы введем словоформу **JEAL**-, то получим следующий результат:

- ✦ jealousy (1)
- ✦ jealous (51)
- ✦ jealousies (11)
- ✦ jealousy (39)
- ✦ loving-jealous (1)

Каждое словоупотребление включает следующий набор признаков. *Контекст* - это связный фрагмент исходного текста, включающий данную словоформу и достаточный для понимания ее смысла и синтаксического окружения. Обычно контекст представляет собой целую строфу, сонет или предложение. *Адрес* данного словоупотребления в корпусе, достаточный для его идентификации при цитировании и предназначенный для человека. Адрес включает в себя:

- 1) название произведения;
- 2) внутренний адрес - названия или номера явно выраженных структурных элементов текста (действие, явление, реплика, ремарка, заголовок) или меток (номер страницы, стиха).

Для драматических произведений внутренний адрес включает в себя: номер действия, номер явления, имя говорящего. Для поэтических произведений внутренний адрес включает номер строфы и стиха. Адрес не всегда обеспечивает однозначную идентификацию текстового фрагмента (может быть несколько реплик одного лица на одной странице, ремарки не имеют имени), но вполне достаточен для цитирования и согласуется с принятой традицией.

Прямая гипертекстовая ссылка указывает на соответствующее место в корпусе текстов.

#	Work	Character	Line	Text
38	Sonnet 57	Shakespeare	785	Being your slave, what should I do but tend Upon the hours and times of your desire? I have no precious time at all to spend, Nor services to do, till you require. Nor dare I chide the world-without-end hour Whilst I, my sovereign, watch the clock for you, Nor think the bitterness of absence sour When you have bid your servant once adieu; Nor dare I question with <i>y jealous</i> thought Where you may be, or your affairs suppose, But, like a sad slave, stay and think of nought Save, where you are how happy you make those. So true a fool is love that in your will, Though you do any thing, he thinks no ill.

Помимо конкорданса словоформ издание включает в себя алфавитный список слов, встречающихся в произведениях поэта и драматурга, также указана их частотность. Стоит отметить удобство в том плане, что такое слово как *play* может встречаться в нескольких формах, например *plays, playing, and played*. Каждое слово присутствует в базе данных наряду с количеством случаев, когда это слово встречается в текстах.

Следует уделить внимание достаточно удобному и эргономичному интерфейсу конкорданса, поиск слова занимает 20-30 секунд. Конкорданс решает многие лингвистические проблемы, уже перечисленные выше, такие как анализ ключевых слов и словосочетаний, частотный анализ лексических единиц. По результатам поиска мы можем не только сказать, сколько раз и в каких произведениях встречается искомая словоформа, но также узнаем, в речи какого персонажа она присутствует, в какой строчке данного произведения. Особо хочется отметить возможность сравнивать сразу два сонета построчно, что позволяет проводить более точный и тщательный анализ. Вряд ли такой результат возможен без конкорданса.

Подводя итоги, скажем, что при составлении конкорданса под массивом текстов подразумеваются как художественные произведения одного или группы авторов, так и тексты деловые, публицистические и пр., т.е. выбор текстового материала всецело определяется целями и задачами, которые ставит перед собой составитель конкорданса. Среди задач словарей типа конкорданса, помимо чисто лингвистических, на которых мы уже останавливались, стоит отметить следующие: отразить все возможные различия и особенности употреблений слов; представить авторское словоупотребление во всей его широте и во всей полноте его языковой специфики; предоставить историко-лексикологическое и лексикографическое описание материалов; уточнить, благодаря анализу иллюстративного материала, т.е. контекстов, лексическое окружение той или иной словоформы, не только ее сочетаемость, но и в некоторых случаях ее значение, особенно это важно при семантизации заимствований и устойчивых словосочетаний; составить частотный словоформуказатель; обобщить материалы источников и расширить возможности использования данных исследуемых текстов в научных исследованиях; проверить распространение лексических единиц по документам разных текстов, учитывая жанровое разнообразие. Говоря о сфере применения лексикографических продуктов, следует, прежде всего, отметить их незаменимость при приведении тех или иных статистических данных.

В первую очередь, это, безусловно, касается определения наиболее частотной лексики (включая все случаи словоупотребления той или иной лексической единицы). Таким образом, конкорданс не только делает возможным осуществление и воплощение в жизнь многочисленных лингвистических и лексикографических задач, но и фактически предоставляет богатейший материал для дальнейших исследований.

Список литературы

1. **Архинов А. В.** Конкорданс [Электронный ресурс]. URL: <http://www.lomonosov-fund.ru/enc/ru/encyclopedia:0127200> (дата обращения: 19.10.2001).
2. **Вафин Б. Д.** Конкорданс как полноправный продукт лексикографической деятельности (на материале памятников старорусской деловой письменности) [Электронный ресурс]. URL: http://www.kls.ksu.ru/boduen/bodart_3.php?id=3&num=4000000 (дата обращения: 18.10.2011).
3. **Герд А. С.** Автоматизация в лексикографии и словари-конкордансы // Филологические науки. 1981. № 1. С. 72-78.
4. **Демченко А. И.** О необходимости конкордансов нового типа // Изв. АН СССР. Сер. Лит. и яз. 1985. Т. 44. № 5. С. 438-444.
5. **Карпова О. М.** Большой оксфордский словарь английского языка: истоки, развитие и современное состояние // Английская лексикография: формирование, развитие, современное состояние: юбил. сб. науч. тр. Иваново, 2006.
6. **Карпова О. М.** Роль писательских словарей в формировании английской национальной лексикографии // Вестник ВГУ. Серия «Лингвистика и межкультурная коммуникация». 2009. № 1. С. 10-14.
7. **Конкорданс по произведениям У. Шекспира** [Электронный ресурс]. URL: <http://www.opensourceshakespeare.com/> (дата обращения: 15.10.2001).
8. **Коробейникова О. В.** Эволюция словарей языка писателей в англоязычной лексикографии [Электронный ресурс]. URL: http://www.rusnauka.com/SND_2007/Philologia/18295.doc.htm (дата обращения: 18.10.2011).
9. **Опыт конкорданса к роману в стихах А. С. Пушкина «Евгений Онегин»:** с приложением текста романа. М.: ЛексЭст, 2003. 592 с.
10. **Поляков А. Е., Бергельсон М. Б., Пильщиков И. А.** Конкорданс к текстам Ломоносова - концепция и реализация [Электронный ресурс]. URL: <http://www.dialog-21.ru/dialog2009/materials/html/61.htm> (дата обращения: 17.10.2011).
11. **Словарь-конкорданс публицистики Ф. М. Достоевского** [Электронный ресурс]. URL: <http://dostoevskii.karelia.ru>
12. **Ступин Л. П.** Нормативность и лексикография // Проблемы лексикографии. Л., 1997. С. 18.
13. **Anonymous Concordance to Shakespeare.** L., 1787.
14. **Ayscough S.** An Index to the Remarkable Passages and Works Made Use of by Shakespeare. L., 1790.
15. **Clark M.** The Complete Concordance to Shakespeare's Poems: an index. L., 1845.
16. **Dyce A.** A Glossary to the Works of William Shakespeare. L., 1902.
17. **Furness H.** A Concordance to Shakespeare's Poems: an index to every word therein contained. Philadelphia, 1875.
18. **Hatch E.** Concordance to the Septuagintal and Other Versions of the Old Testament. NY, 1954.
19. <http://www.antlab.sci.waseda.ac.jp/software.html>
20. <http://www.glossa.fltr.ucl.ac.be/>
21. <http://www-igm.univ-mlv.fr/~unitex/>
22. **Карпова О.** Author Concordances, with Special Reference to Shakespeare // Lexicography: Critical Concepts / ed. by R. R. K. Hartman. L. - NY, 1996.
23. **Nave's Topical Bible** [Электронный ресурс]. URL: <http://www.biblestudytools.com/concordances/naves-topical-bible/> (дата обращения: 16.11.2011).
24. **Schmidt A.** Shakespeare Lexicon. Berlin - L., 1902.
25. **Spevack M.** A Complete and Systematic Concordance to the Works of Shakespeare. Hildesheim, 1968.
26. **Spevack M.** Harvard Concordance to Shakespeare. Harvard Univ. Press, 1973.
27. **Tatlock J., Kennedy A.** A Concordance to the Complete Works of Chaucer. Washington, 1927.

ABOUT WORKING ON CONCORDANCES PROCEDURE

Kseniya Sergeevna Dobrynina

Department of Translation Theory and Practice

Orenburg State Pedagogical University

Ks.thekind@yandex.ru

The article reveals the basic notion "concordance", its origin history, structure, basic functions and tasks. The author opposes dictionaries and concordances according to several scales, pays special attention to the numerous and various as to content and construction principles dictionaries of fiction authors' language and considers in detail W. Shakespeare's open concordance dictionary entry structure.

Key words and phrases: dictionary; concordance; lexicography; language corps; dictionary entry.